Model-based measurement of latent risk in time series with applications

Bijleveld, F.D.; Commandeur, J.J.F.; Gould, P.; Koopman, S.J.

published in
Journal of the Royal Statistical Society. Series A. Statistics in Society
2008

DOI (link to publisher)
10.1111/j.1467-985X.2007.00496.x

document version
Publisher's PDF, also known as Version of record

J. R. Statist. Soc. A (2008)
171, *Part* 1, *pp.* 265–277

Model-based measurement of latent risk in time series with applications

Frits Bijleveld and Jacques Commandeur,

Institute for Road Safety Research, Leidschendam, The Netherlands

Phillip Gould

Monash University, Melbourne, Australia

and Siem Jan Koopman

VU University Amsterdam, The Netherlands

[Received November 2005. Final revision April 2007]

Summary. Risk is at the centre of many policy decisions in companies, governments and other institutions. The risk of road fatalities concerns local governments in planning countermeasures, the risk and severity of counterparty default concerns bank risk managers daily and the risk of infection has actuarial and epidemiological consequences. However, risk cannot be observed directly and it usually varies over time. We introduce a general multivariate time series model for the analysis of risk based on latent processes for the exposure to an event, the risk of that event occurring and the severity of the event. Linear state space methods can be used for the statistical treatment of the model. The new framework is illustrated for time series of insurance claims, credit card purchases and road safety. It is shown that the general methodology can be effectively used in the assessment of risk.

Keywords: Actuarial and financial risk; Dynamic factor analysis; Epidemiology; Kalman filter methods; Maximum likelihood; Road safety; Unobserved components

1. Introduction

In the statistics and econometrics literature the term ‘risk’ can take many meanings. Here we focus on event or operational risk: given a certain level of exposure, what is the expected severity of loss due to certain events? Examples of exposure are the number (or value) of buildings that are owned by a corporate firm or the size of agricultural land with a certain crop. The event can be fire (which is relevant to buildings) and flooding (which is relevant to crops). This definition of risk contrasts with, for example, value at risk where the focus is on the maximum loss with a probability of, say, 1% in a prespecified period. These two approaches of risk can be regarded as complements. Value at risk focuses on the extreme and total risk whereas operational risk is concerned with expected and more specific risk. Government and industry are concerned with a large variety of operational risk in relation to many different events. For example, road safety is of concern to the general public and therefore most governments take an active role in this. Also, insurance companies focus on the risk of a certain claim whereas epidemiologic research is usually concerned with medical risk of infection. There is growing pressure to develop

Address for correspondence: Siem Jan Koopman, Department of Econometrics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands.
E-mail: s.j.koopman@feweb.vu.nl

 0964–1998/08/171265

risk models in a range of fields. International regulations (from the Basel Committee on Bank Supervision) require banks to be able to model and forecast risk. Road safety researchers have considerable pressure from governments to evaluate past safety measures and to forecast future accidents and injuries.

Event or operational risk is generally concerned with

(a) exposure to an event,
(b) the probability that the event will occur and
(c) the severity of the event.

The time series modelling of event risk offers new insights into data and can confirm or reject the validity of constant risk assumptions. There is substantial evidence that simple deterministic models fail to explain the dynamics of risk adequately. Recently several references have examined stochastically time varying structures to model risk in epidemiological applications. For example, Dominici *et al.* (2004) found evidence of time varying risk factors within a generalized additive model framework that was used to determine the interaction between rates of mortality and concentrations of air pollution. Finkenstädt and Grenfell (2000) found evidence of seasonal time variations in a model for measles epidemics. An illustration of modelling incidences of disease on the basis of latent processes was given by Morton and Finkenstädt (2005). In actuarial research, there is a surprising lack of time series models for the risk and severity of insurance claims. Among the few references is de Jong and Boyle (1983), in which Bayesian methods were applied to a state space model which produces stochastically time varying mortality rates. Harvey and Fernandes (1989) also developed a model for insurance claims using latent factors where both the size of claims and the number of claims were modelled. Automobile insurance claims for multiple cohorts were analysed by Ledolter *et al.* (1991), who tested for common latent factors across cohorts. In bank risk management there have been some references examining the use of time varying parameters to model the risk of counterparty default. Allen and Saunders (2003) highlighted the need for dynamic approaches to modelling company default. A time varying logistic model for durations of unemployment was developed for this purpose by Fahrmeir and Wagenpfeil (1996). It assesses the probability of subjects entering or leaving a state of unemployment. The results suggest that there is a need for time variation in model parameters. In road safety research, the framework of Oppe (1989) assumes that exposure follows a logistic S-curve and log-risk evolves deterministically. The *demande routière, des accidents et leur gravité* approach of Gaudry (1984) and Gaudry and Lassarre (2000) uses regression and Box–Jenkins methods for separating the effects of the risk of a crash and exposure. Li and Kim (2000) used cross-sectional methods for this purpose. Levitt and Porter (2001) showed the importance of sample selection in a microeconomic framework for analysing effects of seatbelts and airbags on accident survival rates. Time series approaches are regarded as complementary to cross-sectional methods since they account for serial correlation also and they can be used when only aggregated time series are available.

In this paper we introduce a general multivariate model for event risk analysis that can consider exposure, risk and severity simultaneously. The latent risk time series (LRTS) model can be applied to a range of problems involving event risk and is not specifically limited to particular applications. The LRTS model is general and allows for the stochastic evolution of exposure, risk and severity over time. It extends previous work by treating exposure and severity as an integral part of the risk problem. In existing approaches some or all of these variables (particularly exposure) are treated as known, when in reality they are measured under error and are subject to stochastic variation. The LRTS model has a multivariate structure and therefore correlations between latent processes and errors can be estimated. The multivariate decomposition can include latent factors for trend, seasonal and cyclical dynamics together with regression and intervention effects. It further allows for the forecasting of future exposures, events and losses together with prediction confidence bounds, which are of particular interest to risk managers. Finally, our multivariate framework can also handle data with multiple cohorts.

The statistical framework, including state space forms and estimation methods, is presented in Section 2. The exposure–risk motor vehicle insurance model is the first example of an LRTS analysis and is discussed in Section 3. The exposure–risk–severity model for credit card use is treated in Section 4. The multiple-exposure–single-risk model for bicycle and moped road traffic accidents is presented in Section 5. The empirical illustrations include parameter estimation, signal extraction of latent factors and some discussion of results.

2. The statistical framework

The LRTS model includes latent factors for exposure E_{it}, risk R_{jt} and severity S_{kt} which are associated with the observed variables exposure X_{it}, outcome Y_{jt} and the loss Z_{kt} for subject indices $i=1,\ldots,I$, $j=1,\ldots,J$, $k=1,\ldots,K$ and time index $t=1,\ldots,n$. The basic form of the model is for $I=J=K$ and links the observables with the latent factors via the multiplicative relationships

$$\begin{aligned} X_{it} &= E_{it} \times U_{it}^{(X)}, \\ Y_{it} &= E_{it} \times R_{it} \times U_{it}^{(Y)}, \\ Z_{it} &= E_{it} \times R_{it} \times S_{it} \times U_{it}^{(Z)}, \end{aligned}$$

where $U_{it}^{(a)}$ are random-error terms with unit mean for $i=1,\ldots,I$, $t=1,\ldots,n$ and $a=X,Y,Z$. The exposure variable X_{it} can be the number of vehicle (type i) registrations or distance travelled, the number or value of loans (type i) or population in region i. The outcome variable Y_{it} is typically the number of times that a certain event occurs for a group i such as claims, accidents and successful treatments. The loss variable Z_{it} measures the severity of the outcome such as the dollar value of claims or defaults (type i). The multiplicative error terms reflect that observed variables are measured under uncertainty due to inaccurate reporting and use of proxy variables. It is not needed to set $I=J=K$ because multiple outcomes for only a single exposure variable can occur and multiple types of severity can exist for a single outcome. For example, we can have multiple types of accidents with cars, so $I=1$ and $J>1$.

Variables in logarithms are denoted by the small version of the corresponding capital letter that is used for the original variable, e.g. $e_{it}=\log(E_{it})$. Further, for any t, we denote $v_t=(v_{1t},\ldots,v_{It})'$ where v_{it} represents any variable with two indices i and t and with the first index i used as stacking argument for $i=1,\ldots,I$ and $t=1,\ldots,n$. After taking logarithms (element by element) and stacking variables in vectors, the multiplicative LRTS equations become the linear system

$$\left.\begin{aligned} x_t &= e_t + u_t^{(x)}, \\ y_t &= e_t + r_t + u_t^{(y)}, \\ z_t &= e_t + r_t + s_t + u_t^{(z)}, \end{aligned}\right\} \tag{1}$$

where $u_t^{(a)}$ is a serially independent disturbance vector with zero mean and variance matrix $\Sigma_u^{(aa)}$ for $a=x,y,z$. The disturbances can also be mutually but instantaneously correlated and the corresponding covariance matrix is denoted by $\Sigma_u^{(ab)}$ for $a,b=x,y,z$ and $a\neq b$. In case the dimensions I, J and K do not match, the different series for x, y and z can be distributed generally via

$$\left.\begin{aligned} x_t &= e_t + u_t^{(x)}, \\ y_t &= H_{yx} e_t + r_t + u_t^{(y)}, \\ z_t &= H_{zy}(H_{yx} e_t + r_t) + s_t + u_t^{(z)}, \end{aligned}\right\} \tag{2}$$

where $J \times I$ matrix H_{yx} and $K \times J$ matrix H_{zy} are typically selection matrices consisting of 1s and 0s. It is assumed that the dimensions of observed exposure (proxy) x and latent exposure e, of observed outcome y and latent risk r and of observed loss z and latent severity s match and are equal to I, J and K respectively. It is straightforward to modify system (2) further to account for cases where the dimensions of observed and corresponding latent variable do not match. However, identifiability of the system becomes an issue in such cases whereas this is not the case for system (2) since any latent variable is uniquely linked with an observed variable.

The additive system (1) is the observation equation where log-exposure e_{it}, log-risk r_{it} and log-severity s_{it} are treated as latent factors which can be modelled separately. The latent factors can be specified as vector auto-regressive integrated moving average processes. A more flexible approach is to let these factors depend on a sum of auto-regressive integrated moving average processes and fixed effects as advocated by Harvey (1989), which are known as structural time series models, and Bell (2004), which are known as 'RegComponent' models. For example, latent factor c may partly depend on a trend (long-term) component that is modelled by $\mu_{t+1}^{(c)} = \mu_t^{(c)} + \beta_t^{(c)} + \eta_t^{(c)}$ with $\beta_t^{(c)} = 0$ (the local level model) or with $\beta_{t+1}^{(c)} = \beta_t^{(c)} + \zeta_t^{(c)}$ (the local linear trend model) where $\eta_t^{(c)}$ and $\zeta_t^{(c)}$ are disturbance vectors with zero mean and variance matrices Σ_η^{cc} and Σ_ζ^{cc} respectively, for $c = e, r, s$. The disturbance vectors $\eta_t^{(c)}$ and $\zeta_t^{(c)}$ for latent factor c are mutually independent. However, the contemporaneous covariance matrix between disturbance vectors $\eta_t^{(c)}$ and $\eta_t^{(d)}$, which is denoted by $\Sigma_\eta^{(cd)}$, can be non-zero for $c, d = e, r, s$ and $c \neq d$. This may also apply to $\zeta_t^{(c)}$ and $\zeta_t^{(d)}$.

In case the time series is observed in quarterly or monthly frequencies, the series may be subject to seasonal effects. The latent factor c may then depend on a periodic (seasonal) process that can be modelled by

$$\sum_{j=0}^{p-1} \gamma_{t+1-j}^{(c)} = \omega_t^{(c)}$$

(stochastic seasonal dummy) where $\gamma_t^{(c)}$ is the seasonal effect at time t with seasonal length p for $c = e, r, s$. The disturbance vector $\omega_t^{(c)}$ has similar properties to those of the disturbance vectors $\eta_t^{(c)}$ and $\zeta_t^{(c)}$ but they are mutually independent of each other. Apart from trend and seasonal dynamics, latent factors can be further composed of, possibly stationary, auto-regressive integrated moving average processes.

Regression effects can be added to the latent factor as is common within the frameworks of structural time series models and 'RegComponent' models. Fixed regression effects can also include intervention effects for outlying observations $D_t^{(c)}(\tau; 0)$, level breaks in trend $D_t^{(c)}(\tau; 1)$ and slope breaks in trend $D_t^{(c)}(\tau; 2)$ where $1 < \tau < n$ is a fixed time point at which the intervention occurs for factors $c = e, r, s$. We can formally define the interventions by $D_t^{(c)}(\tau; 0) = 1$, $\Delta D_t^{(c)}(\tau; 1) = 1$ and $\Delta^2 D_t^{(c)}(\tau; 2) = 1$ for $t = \tau$, and all are 0 otherwise, with difference operator $\Delta = 1 - B$ and backshift operator B so that $\Delta y_t = (1 - B) y_t = y_t - y_{t-1}$. An illustration of intervention analysis in this framework was presented by Harvey and Durbin (1986) for a univariate time series of road accidents. A general model-based methodology for identifying interventions from a given time series was developed by de Jong and Penzer (1998).

Components and fixed effects are assumed to be part of the latent factors e_t, r_t and s_t and therefore part of the LRTS system. This implies that a seasonal or an intervention effect in observed exposure also enters the equations for observed outcome and loss. However, the modeller may decide that some effects need to appear exclusively in one equation. We therefore need to introduce the idiosyncratic latent factors $e_t^{(x)}$ and $r_t^{(y)}$ for the observation equations for exposure and outcome respectively. We obtain

$$\left.\begin{aligned} x_t &= e_t + e_t^{(x)} + u_t^{(x)}, \\ y_t &= H_{yx}e_t + r_t + r_t^{(y)} + u_t^{(y)}, \\ z_t &= H_{zy}(H_{yx}e_t + r_t) + s_t + u_t^{(z)}. \end{aligned}\right\} \tag{3}$$

The compositions of the idiosyncratic factors $e_t^{(x)}$ and $r_t^{(y)}$ can be specified in the same way as the factors e_t, r_t and s_t as described above. Some account needs to be taken with respect to the identification of the factors. For example, in the case $I=J=K=1$, a seasonal component can appear in both e_t and $e_t^{(x)}$ but for the remaining two equations only one additional seasonal component is available since only three observed series are given to identify the seasonal effects. This also applies to other effects that are part of the model.

It is well documented (see earlier references in this section) that different linear dynamic processes can be formulated in state space form jointly. The state equation as formulated in Durbin and Koopman (2001) is given by

$$\alpha_{t+1} = T_t\alpha_t + G_t\xi_t, \qquad \xi_t \sim N(0, Q_t), \quad t=1,\ldots,n, \tag{4}$$

where the initial state vector α_1 is specified separately. For example, the local level model that is defined above is obtained from expression (4) by having T_t and G_t as identity matrices and setting $Q_t = \Sigma_\eta^{(cc)}$. Regression effects can also be considered as a part of the state vector. In this framework we define a component as a linear function of the state vector containing latent processes and regression effects. Whereas matrix G_t is typically a known selection matrix, elements of the matrices T_t and Q_t may be unknown, as is apparent from the example above. The unknown elements are collected in the parameter vector ψ and are estimated as described below.

The state space formulation is completed with the observation equation

$$\begin{pmatrix} x_t \\ y_t \\ z_t \end{pmatrix} = \begin{pmatrix} F_I & 0 & 0 & F_I & 0 \\ H_{yx} & F_J & 0 & 0 & F_J \\ H_{zy}H_{yx} & H_{zy} & F_K & 0 & 0 \end{pmatrix}\theta_t + u_t, \qquad \theta_t = W_t\alpha_t, \quad t=1,\ldots,n, \tag{5}$$

with $i \times i$ identity matrix F_i for $i = I, J, K$, signal vector $\theta_t = (e_t', r_t', s_t', e_t^{(x)\prime}, r_t^{(y)\prime})'$ and disturbance vector $u_t = (u_t^{(x)\prime}, u_t^{(y)\prime}, u_t^{(z)\prime})'$. Matrix W_t links the signal θ_t with the state α_t by selecting the appropriate elements of the state vector that contains the components and fixed regression effects that are required for modelling the dependent time series x_t, y_t and z_t.

The state equation (4) and the observation equation (5) define the state space model and enable application of the Kalman filter for filtering the state vector. Filtering refers to the estimation of α_t conditional on observations up to and including time t. Smoothing is similar but the estimation is conditional on all observations (up to and including time n). A related method carries out the computations for smoothing. Both methods also compute mean-squared errors for the estimators. In the case that all disturbances in the model are normally distributed, we obtain minimum mean-squared estimators. When normality is not assumed, they are minimum mean-squared linear estimators. A textbook treatment of state space methods is given by Durbin and Koopman (2001) whereas a non-technical introduction is given by Commandeur and Koopman (2007).

The Kalman filter carries out the prediction error decomposition for a given state space model and a particular value of ψ. This implies that the likelihood function can be evaluated by the Kalman filter for a given ψ. Maximum likelihood estimation of ψ then becomes a standard exercise of numerically maximizing the likelihood function with respect to ψ. In the empirical applications of the LRTS model below, parameters in ψ are limited to the elements of variance matrices such as $\Sigma_\eta^{(cc)}$ given above for the local level model. Regression coefficients can be placed in the state vector. To ensure positive semidefinite variance matrices, a variance matrix is decomposed as $\Sigma_\eta^{(cc)} = M'M$ where M is a symmetric matrix.

3. Case I: a two-dimensional insurance latent risk time series model

The first illustration of the latent risk model concerns insurance policies and claims related to motor vehicle fatalities in Victoria, Australia. We analyse annual time series consisting of the number of vehicle registrations (in thousands; exposure x_t) and the number of claims (in units; outcome y_t) for the years 1950–2001. Registrations are a measure of the total stock of vehicles on Victoria's roads. The two time series are presented in Figs 1(a) and 1(b). The registrations series display an upward smooth trend whereas the fatality claims series have a 'hump' shape, with a peak in the early 1970s. Since registrations have increased monotonically over the past 50 years, the reduction in fatality claims must have been caused by a decrease in risk. Risk reductions have been driven by gradual improvements in vehicle and road design together with increased public awareness. Demographic factors have also been important as a new generation of road users ('baby boomers') began to start driving. Public horror at a road casualty toll of 1034 for Victoria in 1970 led to newspaper declarations of 'war on 1034'. This has been indicative of changing

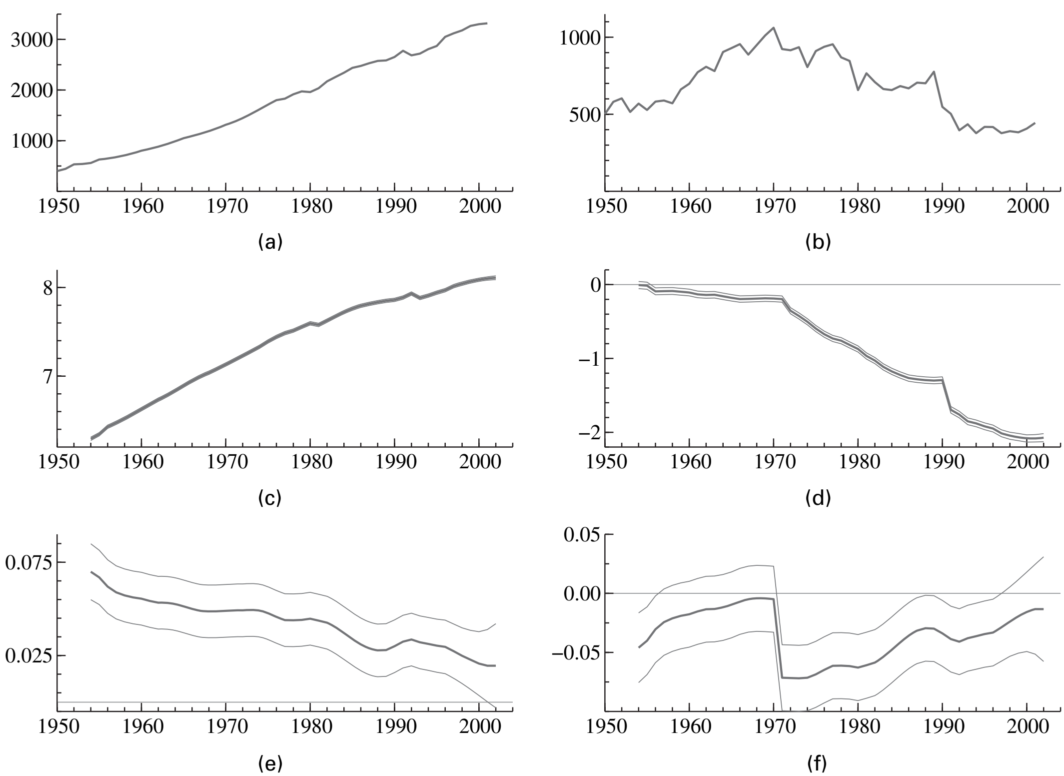

Fig. 1. (a) Time series of registered vehicles (thousands) and (b) crash fatalities, (c), (d) stochastic trends and (e), (f) stochastic slopes including interventions ((a), (c), (e) smooth estimates of exposure; (b), (d), (f) smooth estimates of risk)

attitudes towards road safety. The effects on attitude have proved to be long term. Other important relevant events in the sample are the introduction of seatbelt laws in 1971 and the increased enforcement and mass media advertising campaigns on road safety in the early 1990s.

The policy exposure series x_t and claim outcome series y_t are both univariate (the data are not disaggregated into groups or cohorts). A time series for loss z_t (e.g. the dollar value of pay-outs on claims) is not available and therefore we consider a two-dimensional LRTS model that consists of the first two equations in system (3) with $H_{yx} = 1$ and with dimensions $I = J = 1$. The latent factors e_t and r_t are modelled as local linear trends. The following special events are considered as intervention variables:

(a) in 1970, a publicity campaign was launched to increase public and governmental awareness of road safety issues ('war on 1034');
(b) in 1971, the introduction of seatbelt laws;
(c) in 1980, a change in data collection on vehicle registrations;
(d) in 1990, introduction of advertising and enforcement initiatives aimed at reducing accident risk;
(e) in 1992, another change in data collection on vehicle registrations.

The changes in data collection should only affect exposure and are therefore part of the latent factor e_t whereas the other events should have an effect on risk r_t. Intervention (a) is a long-term effect and therefore is captured by a change in the slope term of risk. The events (b) and (d) are taken as immediate step changes in the level of risk. These interventions are confirmed by applying the methods of de Jong and Penzer (1998) to this data set. Interventions (a), (b) and (d) are assumed to have an impact only on accident risk, as none of the measures are aimed at reducing road use.

Estimates for a selection of parameters are displayed in Table 1. Standard errors are computed but for brevity we do not present them. The estimated (co)variances for trend and slope disturbances for the two latent factors reveal that exposure and risk are perfectly negatively correlated:

$$\Sigma_\eta^{(er)} \Big/ \sqrt{\left(\Sigma_\eta^{(ee)} \Sigma_\eta^{(rr)}\right)} = \Sigma_\zeta^{(er)} \Big/ \sqrt{\left(\Sigma_\zeta^{(ee)} \Sigma_\zeta^{(rr)}\right)} = -1.$$

The perfect negative correlations mean that both exposure and risk factors are subject to the same stochastic shocks that determine their time varying behaviour. This finding is in agreement with most road crash research, which finds a strong negative relationship between risk and exposure. There are various reasons for this relationship, including the fact that roads become more congested as exposure increases, which slows vehicle speeds such that fatal or serious injury accidents are less likely. In developed countries, there has been a period of increased road use and decreasing fatal accident risk over the past 35 years. Over this period, technology and safety awareness have improved, which is also an indirect cause of the negative correlation. The perfect correlation of shocks implies that the components can be interpreted as common factors. Nevertheless, the estimated components are distinct from each other since they are also subject to different interventions.

The estimates of the intervention coefficients are presented in Table 1. The estimated intervention for the expected break in the level of exposure due to a change in the data collection of policies (registrations) is clearly significant for 1992 but less significant for 1980. The level interventions for risk in 1971 (seatbelt laws) and 1990 (advertising initiatives) are very significant. The magnitude of the 1990 intervention is nearly four times greater than that of the seatbelt law that was introduced in 1971. However, the 1971 seatbelt effect may partly be confounded with

Table 1. Parameter estimates for disturbance (co)variances and interventions†

Parameter or intervention	*Description*	*Results for the following cases:*		
		Case I	*Case II*	*Case III*
$\Sigma_\eta^{(ee)}$	Variance trend exposure	0.31×10^{-3}	1.33×10^{-5}	
$\Sigma_\eta^{(rr)}$	Variance trend risk	1.30×10^{-3}	8.91×10^{-5}	
$\Sigma_\eta^{(ss)}$	Variance trend severity		1.18×10^{-5}	
$\Sigma_\eta^{(er)}$	Covariance trend exposure–risk	-0.640×10^{-3}		
$\Sigma_\zeta^{(ee)}$	Variance slope exposure	0.040×10^{-3}	0.0261×10^{-5}	0.27×10^{-4}
$\Sigma_\zeta^{(rr)}$	Variance slope risk	0.130×10^{-3}	0.1590×10^{-5}	20.0×10^{-4}
$\Sigma_\zeta^{(ss)}$	Variance slope severity		0.0014×10^{-5}	
$\Sigma_\zeta^{(er)}$	Covariance slope exposure–risk	-0.070×10^{-3}	-0.0371×10^{-5}	
$\Sigma_\zeta^{(es)}$	Covariance slope exposure–severity		-0.0058×10^{-5}	
$\Sigma_\zeta^{(rs)}$	Covariance slope risk–severity		0.0056×10^{-5}	
$\Sigma_\omega^{(yy)}$	Variance seasonal outcome		220×10^{-5}	
$\Sigma_\omega^{(zz)}$	Variance seasonal loss		181×10^{-5}	
$\Sigma_\omega^{(yz)}$	Covariance seasonal outcome–loss		194×10^{-5}	
$\Sigma_u^{(xx)}$	Variance disturbance exposure	0.16×10^{-3}	0.31×10^{-5}	9.70×10^{-4}
$\Sigma_u^{(yy)}$	Variance disturbance outcome	4.21×10^{-3}	1.07×10^{-5}	0.47×10^{-4}
$\Sigma_u^{(zz)}$	Variance disturbance loss		4.13×10^{-5}	
$D_t^{(r)}(1970; 2)$	'War on 1034'	−0.079‡		
$D_t^{(r)}(1971; 1)$	Seatbelt law introduction	−0.108‡		
$D_t^{(e)}(1980; 1)$	Data collection change	−0.086§		
$D_t^{(r)}(1990; 1)$	Advertising initiative	−0.376‡		
$D_t^{(e)}(1992; 1)$	Data collection change	−0.066‡		
$D_t^{(x)}(2002.1; 1)$	Data collection change		0.062‡	
$D_t^{(y)}(2002.1; 1)$	Data collection change		0.066§	
$D_t^{(s)}(2002.1; 1)$	Data collection change		0.083‡	
$D_t^{(e)}(1991; 1)$	Free travel pass introduction			−0.180§
$D_t^{(r)}(2000; 1)$	Start of law on mopeds on main roads			−0.310‡

†The last three columns are for the three models that are described in the sections for cases I, II and III. We consider observed time series for exposure (x), outcome (y) and loss (z) which are modelled by expression (1) with the latent factors for exposure (e), risk (r) and severity (s). Model (1) is considered for case I (without loss and severity), case II and case III (without loss and severity). The latent factors are in all cases modelled by local linear trend models and in case II together with a stochastic seasonal dummy component. The covariance matrix $\Sigma_\eta^{(cd)}$ is for the trend component, $\Sigma_\zeta^{(cd)}$ is for the slope component of the trend and $\Sigma_\omega^{(cd)}$ is for the stochastic seasonal dummy component, for $c, d = e, r, s$ in model (1). The covariance matrix $\Sigma_u^{(ab)}$ is for the disturbance vector $u_t^{(a)}$, for $a, b = x, y, z$ in model (1). The intervention effect $D_t^{(c)}(\tau; 1)$ is for a level break in the trend component and $D_t^{(c)}(\tau; 2)$ is for a slope break in the trend with $c = e, r, s$ and time point τ at which the intervention occurs.
‡Significance at the 95% level.
§Significance at the 90% level.

the highly significant 'war on 1034' effect on the slope of the log-risk. This estimated effect of −0.079 implies that each year a reduction of 0.079 is achieved in log-risk. The combined effects of 1970 and 1971 have therefore more effect than the advertising campaign in 1990. Since the different events occur shortly after the beginning of the 1970s, it is difficult to disentangle those effects.

Fig. 1 presents the estimated level and slope components of exposure and risk (in logarithms). The estimated components are subject to both random shocks and interventions. The salient features of the analysis are the increasing exposure with a significant slope term throughout the sample, and the decreasing risk with a significant negative slope term that is mainly caused by the publicity intervention. Risk displays relatively more stochastic variation in both the estimated level and the slope terms. Apart from the intervention shocks, level and slope components of risk are perfectly and negatively correlated with level and slope components of exposure respectively. The estimated slopes of risk and exposure are of opposite sign but both evolve towards 0. This suggests a long-term flattening of risk and exposure, which is evident in the data. The level terms are also perfectly and negatively correlated. As exposure increases around its slope, risk decreases. Exposure evolves relatively smoothly, with the slope term driving much of the variation.

4. Case II: a three-dimensional credit card latent risk time series model

In this section we study the developments in the usage of credit cards in Australia. The data set consists of monthly observations, from May 1994 through to August 2004 (124 observations), with the number of credit card accounts (exposure x_t), the number of purchases made by credit cards (outcome y_t) and the total dollar value of purchases by credit cards (loss z_t), as presented in Figs 2(a)–2(c). The analysis is crucial for marketing credit cards but is also of concern to bank risk managers who have an interest in Australian consumers' reliance on credit card debt. Since the observed time series for x_t, y_t and z_t have (rapid) increasing patterns, we model the latent factors e_t, r_t and s_t as local linear trends. The monthly series y_t and z_t also have seasonal fluctuations around the trend due to changing consumer behaviour within the year due to, for example, Christmas and Easter. The seasonal factors should not necessarily affect risk and severity and therefore we adopt model (3) with $r_t^{(y)}$ and s_t as stochastic seasonal dummy processes. The data are in nominal terms so severity includes inflationary effects. Furthermore, we examine the event of January 2002 when the Reserve Bank of Australia started to include credit card accounts from commercial banks and other financial institutions in the sample. The inclusion of data from other credit card issuers means that the number of credit cards has increased but the unobserved factors risk and severity may also change since the new issuers in the sample of credit card users may represent customers with different spending patterns. The change in the composition of the sample in January 2002 is permanent and therefore level interventions for this month are appropriate and are included for the latent trend factors e_t, r_t and s_t.

The parameter estimates are given in Table 1. The variance matrices for trend and observation noises are taken as diagonal. This is strongly supported by the fact that maximum likelihood estimation produces almost equal likelihood values for models with and without this restriction. The estimated variances of the seasonal disturbances are relatively large compared with the observation noise. Further, the estimate for the seasonal covariance $\Sigma_\omega^{(yz)}$ implies a high correlation and it may therefore be sufficient to consider model (2) with the inclusion of a seasonal component for r_t only. The estimated trends that are presented in Fig. 2 are smooth and their slopes are varying over time. The log-risk growth is decreasing from 1999 onwards whereas the severity growth is more constant over time. Exposure growth is hump shaped. The three intervention estimates are highly significant and are added to the estimated trends in Fig. 2 although they are part of the observation equations. Although the risk factor is significantly affected by the intervention for the change in composition of the survey, the severity of credit card purchases increased the most. It can therefore be concluded that the new account holders in the survey from January 2002 onwards are making more expensive purchases with their

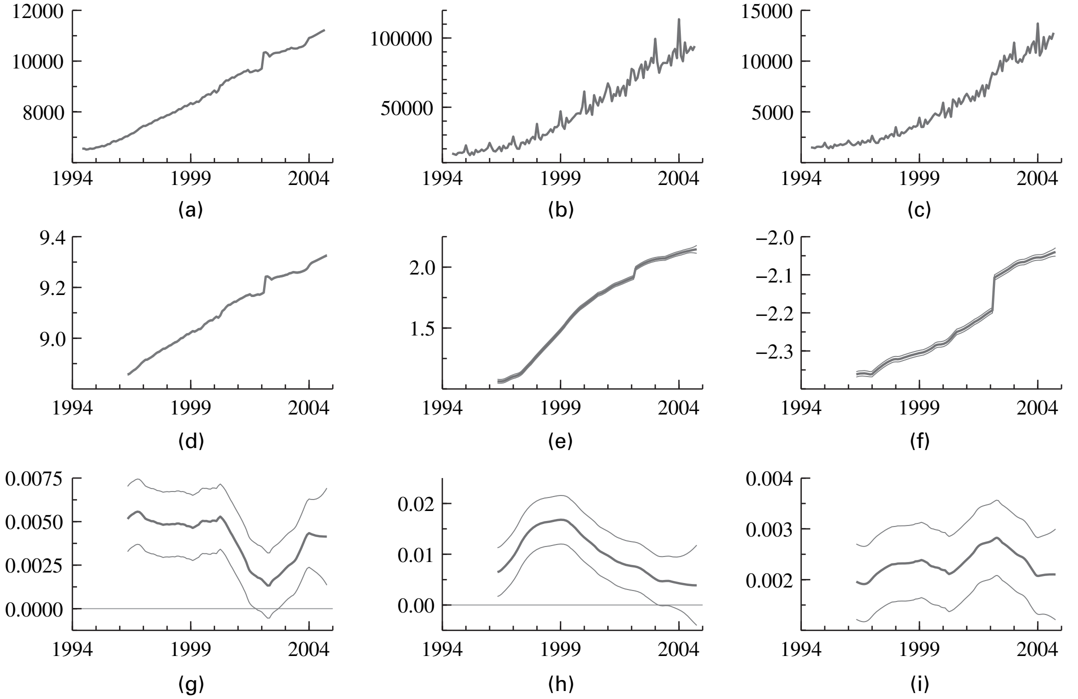

Fig. 2. Monthly time series related to the credit cards data from Australia: (a) number of cards x_t, (b) number of purchases y_t and (c) their value z_t, (d), (e), (f) stochastic trends and (g), (h), (i) stochastic slopes (intervention estimates are added to the trends) ((a), (d), (g) smooth estimates of exposure; (b), (e), (h) smooth estimates of risk; (c), (f), (i) smooth estimates of severity)

credit cards. The new customers have had a smaller effect on the risk (intensity) of making a purchase.

5. Case III: a multiple-exposure latent risk time series model

The yearly numbers of people who are killed and seriously injured in collisions between mopeds and bicycles in the Netherlands are closely watched since they involve mostly young people. Further, mopeds and bicycles are widely and intensively used in the Netherlands. An official study was carried out to investigate the risk of this category of accidents where people are killed or seriously injured. For this, a data set has been constructed with two exposure variables ($I=2$) and one outcome variable ($J=1$). The two exposure variables consist of the numbers of kilometres driven by mopeds and by bicycles. The outcome variable is the yearly number of accidents where the primary collision partners are one moped user and one bicycle user, and where the victims are either killed or hospitalized. The yearly observations range from 1985 to 2003. Given the short sample, the model that was used was parsimonious to preserve a sufficient number of degrees of freedom.

The three time series are presented in Figs 3(a)–3(c). For the two exposure series, the 95% confidence intervals are also presented. These are based on the published survey error variances. The number of kilometres driven by bicycles are subject to stepwise increases in the late 1980s and in 1994 whereas those by mopeds show a gradual decrease over the years. The increase in 1994 for the bicycle-kilometres driven may be explained by an extension of the sample to children under 12 years of age. The decrease in the 95% confidence intervals for the two exposure series from

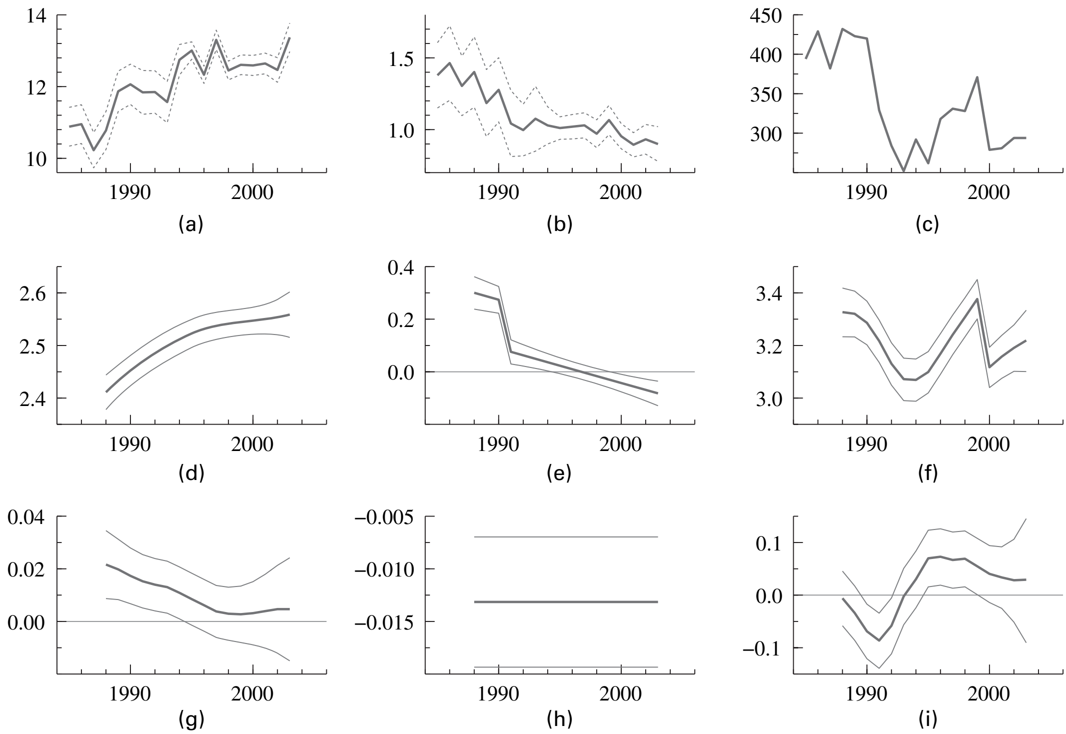

Fig. 3. Yearly time series of traffic volume (in billion kilometres) of (a) bicycles, (b) mopeds together with (c) counts of accidents between them in the Netherlands, (d), (e), (f) stochastic trends and (g), (h), (i) stochastic slopes ((a), (d), (g) smooth estimates of exposure for bicycles; (b), (e), (h) smooth estimates of exposure for mopeds; (c), (f), (i) smooth estimates of risk)

1994 onwards is due to the increase in the survey sample size by a factor of 2. The yearly numbers of accidents show stepwise decreases in 1991 and in 2000. It is expected that the decrease in 1991 coincides with the introduction of a free travel pass for students (typically between 17 and 21 years of age). The travel pass gave free access to the national and local public transport systems (mainly buses and trains). The usage of the free travel pass became increasingly restricted over the years from 1995 onwards. This may partly explain the slow increase of accidents involving death or serious injury in the late 1990s. It is reasonable to argue that the decrease in 2000 may have been caused in part by the introduction of a law that moved all mopeds from the special bicycle roads (or tracks) to the main roads that are in use by other motorized vehicles (motor cars, trucks etc.). This law applies only to situations where special bicycle roads or tracks exist and where the traffic conditions are sufficiently safe. Therefore many exceptions to this law exist and the 'mopeds on the roadway' law can only partly explain the 2000 drop.

The first two equations of the LRTS model (2) are considered with $I=2$, $J=1$ and $H_{yx}=(1\ 1)$. The two latent factors in e_t and the latent factor r_t are modelled as local linear trends. All variance matrices are diagonal. The two variances of $u_t^{(x)}$ depend on a parameter plus a known time varying value that is implied by the different precisions of the surveys. This also applies to the variance of $u_t^{(y)}$ but the time varying value is implied by the normal approximation of the Poisson counts of accidents. The estimated parameters are reported in Table 1. Given the short time span of the sample, the time variations in the level and slope components are limited. The variances of the level disturbances are estimated as 0. In the case of kilometres driven by mopeds, the variation in slope is also estimated as 0 and therefore we obtain a fixed time trend that is only interrupted by the estimated intervention in 1991. The constant variance of the observation noise for moped volume is estimated as 0 so the random noise is due only to the variation in the different sample sizes over the years.

Two significant intervention estimates are reported in Table 1. The first estimate is for the effect of the variable representing the introduction of the free travel pass in 1991 on kilometres driven by mopeds. The second estimate is for the variable representing the effect of the law of mopeds on the roadway on risk. The extension of the sample for bicycle volume with children under 12 years of age did not affect the analysis. We also have experimented with other possible interventions but their inclusion had little or no effect on the value of the likelihood function. The estimated smooth trends for exposure and risk are displayed in Fig. 3. Risk is decreasing until the early 1990s but has been increasing since 1993. The estimated slope pattern for risk may be explained by the popularity of light mopeds for which it is not obligatory to wear a crash-helmet. It is evident that accidents are likely to be more severe when the moped drivers concerned do not wear helmets. This may explain the increasing trend in accidents in which people are killed or seriously injured.

6. Conclusions

In this paper we propose an LRTS model for measuring event risk. The multivariate modelling framework includes latent dynamic factors for exposure, risk and severity. The multivariate nature of the model means that common factors can be identified through the correlation structure of latent dynamic processes. The magnitude and sign of correlations may provide interesting interpretations for researchers. The stochastic trend and seasonal factors are time varying by nature and arbitrary recalibrations of model parameters are not needed. This is an advantage that is inherent in the unobserved components time series modelling approach.

The application to credit cards data showed that stochastic variation is important in measuring the risk and severity of credit card purchases. For the car insurance data, stochastic variation seems less important. It appears that structural breaks explain most of the changes in risk and exposure over the past 50 years. The illustration of accidents between mopeds and bicycles has shown that the model can also include multiple categories of exposure variables. When more data are available, more detailed categories of exposure, risk and severity can be considered. For example, different risk factors can be included for males or females, different age groups and different regions. Future research is directed at extending the modelling framework further for handling multiple categories or panel (longitudinal) structures in data.

References

Allen, L. and Saunders, A. (2003) A survey of cyclical effects in credit risk measurement models. *Technical Report 126*. Bank for International Settlements, Basle.

Bell, W. R. (2004) On RegComponent time series models and their applications. In *State Space and Unobserved Components Models: Theory and Applications* (eds A. C. Harvey, S. J. Koopman and N. Shephard). Cambridge: Cambridge University Press.

Commandeur, J. J. F. and Koopman, S. J. (2007) *An Introduction to State Space Time Series Analysis*. Oxford: Oxford University Press.

Dominici, F., McDermott, A. and Hastie, T. (2004) Improved semiparametric time series models of air pollution and mortality. *J. Am. Statist. Ass.*, **99**, 938–948.

Durbin, J. and Koopman, S. J. (2001) *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.

Fahrmeir, L. and Wagenpfeil, S. (1996) Smoothing hazard functions and time-varying effects in discrete duration and competing risk models. *J. Am. Statist. Ass.*, **91**, 1584–1594.

Finkenstädt, B. F. and Grenfell, B. T. (2000) Time series modelling of childhood diseases: a dynamical systems approach. *Appl. Statist.*, **49**, 187–205.
Gaudry, M. (1984) DRAG, un modèle de la demande routière des accidents et leur gravité appliqué au Québec de 1956–1986. *Technical Report CRT-359*. Centre de Recherche sur les Transports, Département de Sciences Économiques, Université de Montréal, Montreal.
Gaudry, M. and Lassarre, S. (eds) (2000) *Structural Road Accident Models: the International DRAG Family*. Oxford: Elsevier Science.
Harvey, A. C. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.
Harvey, A. C. and Durbin, J. (1986) The effects of seat belt legislation on British road casualties: a case study in structural time series modelling (with discussion). *J. R. Statist. Soc.* A, **149**, 187–227.
Harvey, A. and Fernandes, C. (1989) Time series models for insurance claims. *J. Inst. Act.*, **116**, 513–528.
de Jong, P. and Boyle, P. (1983) Monitoring mortality: a state-space approach. *J. Econometr.*, **23**, 131–146.
de Jong, P. and Penzer, J. (1998) Diagnosing shocks in time series. *J. Am. Statist. Ass.*, **93**, 796–806.
Ledolter, J., Klugman, S. and Lee, C. (1991) Credibility models with time-varying trend components. *ASTIN Bull.*, **21**, 73–91.
Levitt, S. D. and Porter, J. (2001) Sample selection in the estimation of air bag and seat belt effectiveness. *Rev. Econ. Statist.*, **83**, 603–615.
Li, L. and Kim, K. (2000) Estimating driver crash risks based on the extended Bradley–Terry model: an induced exposure method. *J. R. Statist. Soc.* A, **163**, 227–240.
Morton, A. and Finkenstädt, B. F. (2005) Discrete time modelling of disease incidence time series by using Markov chain Monte Carlo methods. *Appl. Statist.*, **54**, 575–594.
Oppe, S. (1989) Macroscopic models for traffic and traffic safety. *Accid. Anal. Prevn*, **21**, 225–232.